Pseudo-Bayesian DIP Denoising as a Preprocessing Step for Kinetic Modelling in Dynamic PET

Alexander C. Whitehead *Student Member, IEEE*, Kjell Erlandsson, Ander Biguri, Scott D. Wollenweber *Senior Member IEEE*, Jamie R. McClelland and Kris Thielemans *Senior Member, IEEE*

***Abstract*—Noise (among other artefacts) could be considered to be the bane of PET. Many methods have been proposed to alleviate the worst annoyances of noise, however, not many take into account the temporal nature of dynamically acquired PET. Here, we propose an adaption of a method, which has seen increasing attention in more traditional imaging denoising circles. Deep Image Prior exploits the initialisation of a carefully designed neural network, so as to treat it as a bank of custom filters, which are to be trained and used afresh on each new image, independently. Deep Image Prior has seen adaptation to PET previously (including dynamic PET), however, many of these adaptations do not take into account the large memory requirements of the method. Additionally, most previous work does not address the main weakness of the Deep Image Prior, its stopping criteria. Here, we propose a method which is both memory efficient, and includes a smoothing regularisation. In addition, we provide uncertainty estimates by incorporating a Bayesian approximation (using dropout), and prototype a training scheme by which the model is fit on all data simultaneously. The denoised images are then used as input for kinetic modelling. To evaluate the method, dynamic XCAT simulations have been produced, with a field of view of the lung and liver. The results of the new methods (along with total variation and the old Deep Image Prior) have been compared by; a visual analysis, SSIM, and K_i values. Results indicate that the new methods potentially outperform the old methods, without increasing computation time, while reducing system requirements.**

Manuscript received November 19, 2022.

This research was supported by GE Healthcare, the NIHR UCLH Biomedical Research Centre and the UCL EPSRC Centre for Doctoral Training in Intelligent, Integrated Imaging in Healthcare (i4health) grant (EP/L016478/1).

The software used was partly produced by the Computational Collaborative Project on Synergistic Biomedical Imaging, CCP SyneRBI, UK EPSRC grant (EP/T026693/1).

Jamie R. McClelland is supported by a Cancer Research UK Centres Network Accelerator Award grant (A21993) to the ART-NET consortium and a CRUK Multi-disciplinary grant (CRC 521).

Alexander C. Whitehead was with the Institute of Nuclear Medicine, University College London, London, UK and the Centre for Medical Image Computing, University College London, London, UK. He is now with the Department of Computer Science, University College London, London, UK (contact: alexander.whitehead.18@ucl.ac.uk).

Kjell Erlandsson and Kris Thielemans are with the Institute of Nuclear Medicine, University College London, London, UK.

Ander Biguri was with the Institute of Nuclear Medicine, University College London, London, UK. He is now with the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK

Scott D. Wollenweber is with Molecular Imaging and Computed Tomography Engineering, GE Healthcare, Waukesha, USA

Jamie R. McClelland and Kris Thielemans are with the Centre for Medical Image Computing, University College London, London, UK

I. INTRODUCTION

MOST Machine Learning or Neural Network (NN) based methods, rely upon a workflow where a model is designed, trained, validated, and deployed [1]. However, in the domain of image denoising, the Deep Image Prior (DIP) method has received attention as training and inference are performed independently on each new image [2].

For PET, there have been a number of adaptations of DIP. [3] used a U-Net with relatively high count/low motion brain scans [4]. In [5] DIP is extended to 4D dynamic PET. To do this, multiple output branches are grafted onto the NN, one for each dynamic time point. [6] uses the original or static PET acquisition as input to the NN, rather than noise. [7] represents a more recent extension, where multiple NNs are used simultaneously.

This work seeks to extend or simplify previous work, in order to denoise 4D dynamic PET data.

II. METHODS

A. Network Design and Execution

Firstly, we reduced GPU memory requirements. Because DIP requires training at inference, the full amount of memory is needed every time the method is used. To aid in clinical adoption, a hard limit of $8.0\,\mathrm{GB}$ GPU memory was imposed. Secondly, a more robust stopping criteria is required. To aid in achieving this, as well as to address weaknesses of the original DIP, regularisation must be added, so as to stop the NN fitting to noise, similar to [8]. Finally, because the output from this method is to be used in a further kinetic model fitting, it may provide improved

978-1-6654-8872-3/22/$31.00 ©2022 IEEE

results to have a metric of the uncertainty in the denoised images. we used dropout as in [9] to approximate (more expensive) Bayesian inference. Furthermore, this work uses PET data with a Field of View (FOV) of the lung and liver, whereas most previous work uses a FOV of the head.

The NN used was a modified U-Net [4], with seven down/upsampling stages. Each down/upsampling stage consisted of two convolutional layers (with two, four, eight, 16, 32, 64, or 128 channels, depending on depth). Followed by, either a split strided convolution and max-pooling layer (with the result concatenated), or a tri-linear upsampling layer. Edge padding, group normalisation [10], MISH activation [11], and spatial dropout, were used with every convolutional layer. Data was edge padded to the nearest power of two, and the input data had Gaussian noise summed to it. Both input and label data were standardised. Mean Squared Error and Total Variation (TV) were used as the loss function. AdamW [12] was used as optimiser. Training continued for all methods until the gradient of the loss function, over a window of previous results, reduced below a threshold. Parameters were tuned using a grid search.

Two training regimes were explored, one where each time point was treated independently, and another, where the model weights were saved and then independently updated on each time point (the mean of the new models weights was taken for the next iteration).

B. PET Acquisition Simulation and Image Reconstruction

A series of dynamic scans, following the clinical Dynamic Whole Body-PET protocol, were generated using the XCAT phantom [13]. Patient-derived kinetic parameters were assigned to; 64 tissues, three tumours of $1.0\,\mathrm{cm}$ diameter in the left lung, and three tumours of $2.5\,\mathrm{cm}$, $2.0\,\mathrm{cm}$ and $1.0\,\mathrm{cm}$ diameter in the liver. An input function for 18F-FDG, taken from [14], was used to simulate Time Activity Curves (TACs) to create dynamic images.

PET acquisitions were simulated (and reconstructed) using STIR [15] through SIRF [16]. Non-Time-of-Flight sinogram data were simulated, using resolution modelling (using a $6.0\,\mathrm{cm}$ FWHM Gaussian filter). Randoms and scatter were not included. Poisson noise was added.

Finally, all data sets were reconstructed using 10 iterations with 17 subsets of OSEM [17].

C. Kinetic Modelling

Indirect Patlak estimation was used to generate K_i and intercept images [18]. Volumes of Interest (VOIs) were defined for the three lesions and two background regions (lung and liver), and the mean values of K_i and V_d were calculated in each VOI. The uncertainties of the parameters were estimated as follows: Normally distributed noise was added to the dynamic images, with standard deviation given by the DIP-uncertainty, and Patlak analysis was performed. The procedure was repeated for 10 noise realisations, and the standard deviation of the K_i and V_d parameters were calculated.

TABLE I

COMPARISON OF $\mathrm{SUV}_{\mathrm{MAX}}$ AND $\mathrm{SUV}_{\mathrm{PEAK}}$, FOR; THE GROUND TRUTH, THE ORIGINAL NOISY DATA, AND THIS DATA DENOISED USING, TV, THE IMPLEMENTATION OF DIP FROM [3], AND OUR NEW IMPLEMENTATION OF DIP, TRAINED SEQUENTIALLY AND COMBINED (TAKEN FOR THE LAST TIME POINT OF THE LUNG FOV).

SUV	Max	Peak
Ground Truth	12.3	9.53
Noisy	21.5	5.99
TV	3.28	6.24
Original DIP	3.90	6.93
New DIP Sequential	9.19	8.02
New DIP Combined	9.27	8.21

D. Evaluation

In addition to the denoising performed above, data were also denoised using TV, and the DIP method presented in [3].

Comparisons used included: A visual analysis, Structural Similarity Index Measure (SSIM) to the ground truth [19], K_i values, a TAC through a lesion, a profile over a lesion, and $\mathrm{SUV}_{\mathrm{max}}$ and $\mathrm{SUV}_{\mathrm{peak}}$ (defined following EANM guidelines [20]).

III. RESULTS

A visual comparison of the reconstructed images (see Fig. 1), shows that both of the new DIP methods perform comparably, if not for a slight reduction in noise in the combined case. Whereas, the TV and original DIP implementations appear to have struggled with over smoothing, reducing the contrast of the lesions, and introducing some edge artefacts. The uncertainty of the combined method can be seen reduced compared to the sequential method.

A comparison of K_i values across multiple lesions (see Fig. 2), shows that the new DIP combined method most often estimates the greatest magnitude of K_i value, which is usually closest to the ground truth. The new DIP sequential is slightly less accurate (also with greater

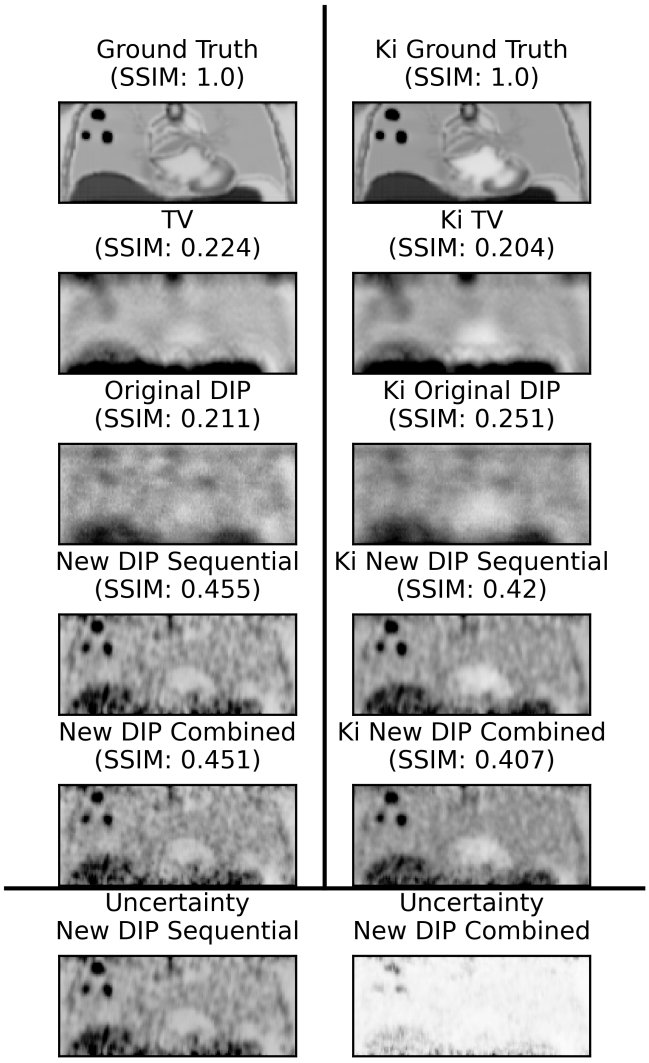

Fig. 1. First column contains, a visual analysis between the ground truth and denoised results (taken for the last time point, plus SSIM to the ground truth), and the second column contains, the K_i results (all voxels in a coronal view) of a Patlak reconstruction of all time points (plus SSIM to the ground truth), for; the ground truth, and data denoised using, TV, the implementation of DIP from [3], and our new implementation of DIP, trained sequentially and combined (taken for the lung FOV). Last row contains, uncertainty volumes, for; the data denoised using our new implementation of DIP, trained sequentially and combined (taken for the last time point of the lung FOV). Colour map ranges are consistent for all images in each section.

uncertainty), however, is more accurate than the TV and original DIP implementations (which consistently significantly underestimate K_i).

The overall shape of the TAC (see Fig. 3) for the new DIP combined method appears, most similar to the ground truth, however, with a slight reduction in quantification. The sequential method is less accurate, but still more so than both the original DIP and TV methods. There is significant variation in the noise TAC, somewhat masked

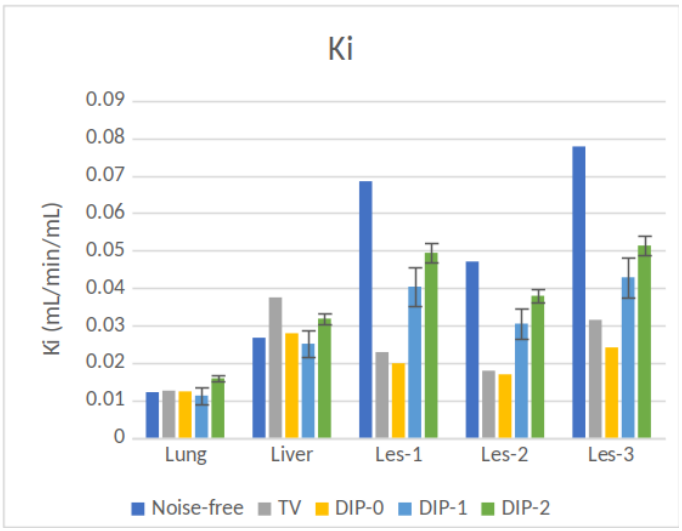

Fig. 2. K_i results (single voxel) of a Patlak reconstruction of all time points, plus uncertainty where applicable, for; the ground truth, and data denoised using, TV, the implementation of DIP from [3], and our new implementation of DIP, trained sequentially and combined.

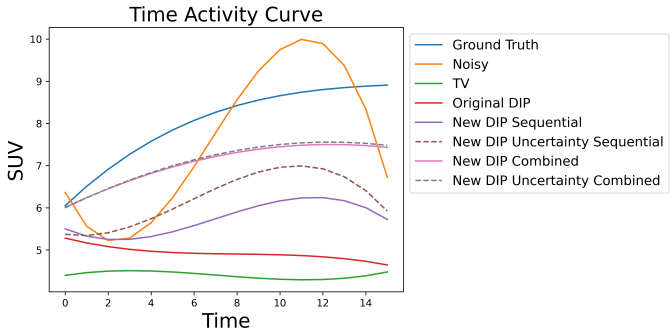

Fig. 3. A TAC through a lesion, fit as a third order polynomial regression, with weighting using uncertainty (where available), for; the ground truth, the original noisy data, and this data denoised using, TV, the implementation of DIP from [3], and our new implementation of DIP, trained sequentially and combined, both with and without uncertainty (taken for the lung FOV).

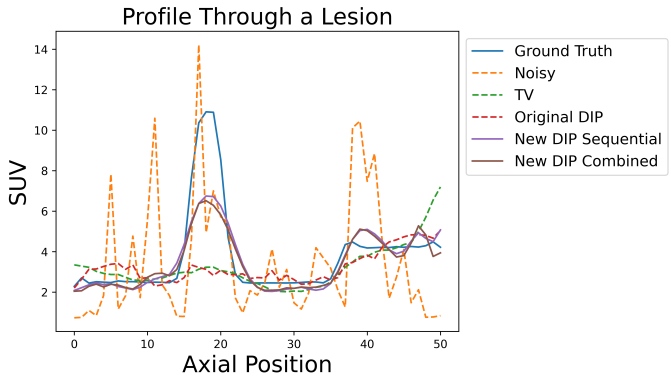

Fig. 4. A profile through a lesion, in the SI direction, for; the ground truth, the original noisy data, and this data denoised using, TV, the implementation of DIP from [3], and our new implementation of DIP, trained sequentially and combined (taken for the last time point of the lung FOV).

by the regression, however, its shape is still least like the ground truth. Adding uncertainty appears to have improved the TAC of the new DIP sequential method, however, the uncertainty of the combined method is less, and as such, the inclusion of uncertainty has not affected results significantly.

The peak of the profile (see Fig. 4) for both new DIP methods is comparable, and greater than both the original DIP and TV methods. The peak of the noise profile is greater than all other methods, including the ground truth, however, this is not necessarily beneficial, as can be seen by the rest of the profile not closely following the ground truth (it undulates unpredictably). The profile for both new DIP methods are significantly smoother, and more closely follow the ground truth.

SUV (and SSIM) results confirm the above (see TABLE I).

IV. DISCUSSION AND CONCLUSION

Evaluation indicated that the new DIP method, particularly when trained combined, provided images with less noise and more quantitative accuracy than other methods. The combined method had lower uncertainty.

Results presented here were obtained on a single bed position. Initial evaluation on a bed position, centred on the liver, indicated that parameter fine-tuning, depending on the distribution and count level, will be beneficial. Evaluation with patient data will follow.

The uncertainty estimates produced by the NN need to be validated by comparison with results obtained from repeated noise realisations.

REFERENCES

[1] B. Krose *et al.*, *An introduction to neural networks*. 2011.

[2] D. Ulyanov *et al.*, "Deep image prior," *International Journal of Computer Vision*, vol. 128, no. 7, pp. 9446–9454, Jul. 2018.

[3] K. Gong *et al.*, "PET Image Reconstruction Using Deep Image Prior," *IEEE Transactions on Medical Imaging*, vol. 38, no. 7, pp. 1655–1665, Jul. 2019.

[4] W. Weng *et al.*, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *IEEE Access*, vol. 9, pp. 16 591–16 603, May 2015.

[5] F. Hashimoto *et al.*, "4D deep image prior: dynamic PET image denoising using an unsupervised four-dimensional branch convolutional neural network," *Physics in Medicine & Biology*, vol. 66, no. 1, p. 015 006, Jan. 2021.

[6] F. Hashimoto *et al.*, "Dynamic PET Image Denoising Using Deep Convolutional Neural Networks Without Prior Training Datasets," *IEEE Access*, vol. 7, pp. 96 594–96 603, 2019.

[7] C. H. Yang *et al.*, "Simultaneous Denoising of Dynamic PET Images Based on Deep Image Prior," *Journal of Digital Imaging 2022*, pp. 1–12, Mar. 2022.

[8] J. Liu *et al.*, "Image Restoration Using Total Variation Regularized Deep Image Prior," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2019-May, pp. 7715–7719, May 2019.

[9] Y. Gal *et al.*, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," *33rd International Conference on Machine Learning, ICML 2016*, vol. 3, pp. 1651–1660, Jun. 2015.

[10] Y. Wu *et al.*, "Group Normalization," *International Journal of Computer Vision*, vol. 128, no. 3, pp. 742–755, Mar. 2018.

[11] D. Misra, "Mish: A Self Regularized Non-Monotonic Activation Function," in *British Machine Vision Conference*, Aug. 2020.

[12] I. Loshchilov *et al.*, "Decoupled Weight Decay Regularization," *7th International Conference on Learning Representations, ICLR 2019*, Nov. 2017.

[13] W. P. Segars *et al.*, "4D XCAT phantom for multimodality imaging research," eng, *Medical Physics*, vol. 37, no. 9, pp. 4902–4915, Sep. 2010.

[14] J. W. Långsjö *et al.*, "Effects of subanesthetic ketamine on regional cerebral glucose metabolism in humans," eng, *Anesthesiology*, vol. 100, no. 5, pp. 1065–1071, May 2004.

[15] K. Thielemans *et al.*, "STIR: software for tomographic image reconstruction release 2," *Physics in Medicine and Biology*, vol. 57, no. 4, pp. 867–883, Feb. 2012.

[16] E. Ovtchinnikov *et al.*, "SIRF: Synergistic Image Reconstruction Framework," in *IEEE NSS-MIC*, IEEE, Oct. 2017, pp. 1–3.

[17] H. M. Hudson *et al.*, "Accelerated Image Reconstruction Using Ordered Subsets of Projection Data," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[18] C. S. Patlak *et al.*, "Graphical evaluation of blood-to-brain transfer constant from multiple-time uptake data," *Journal of Cerebral Blood Flow & Metabolism*, vol. 3, pp. 1–7+, 1983.

[19] Z. Wang *et al.*, "Mean squared error: Lot it or leave it? A new look at signal fidelity measures,"

IEEE Signal Processing Magazine, vol. 26, no. 1, pp. 98–117, 2009.

[20] R. Boellaard *et al.*, "FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 42, no. 2, pp. 328–354, 2015.